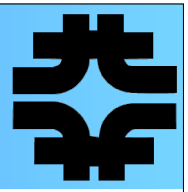


US-CMS Software and Computing Project and Networking

Ian M. Fisk

September 15, 2004

Introduction to CMS

CMS is one of two multi-purpose detectors being build for the LHC

SUPERCONDUCTING COIL

Basic Choices:

Strong, large B field (4T)

All Si tracking (L)

Best possible ECAL dE/E

Robust Muon - yoke

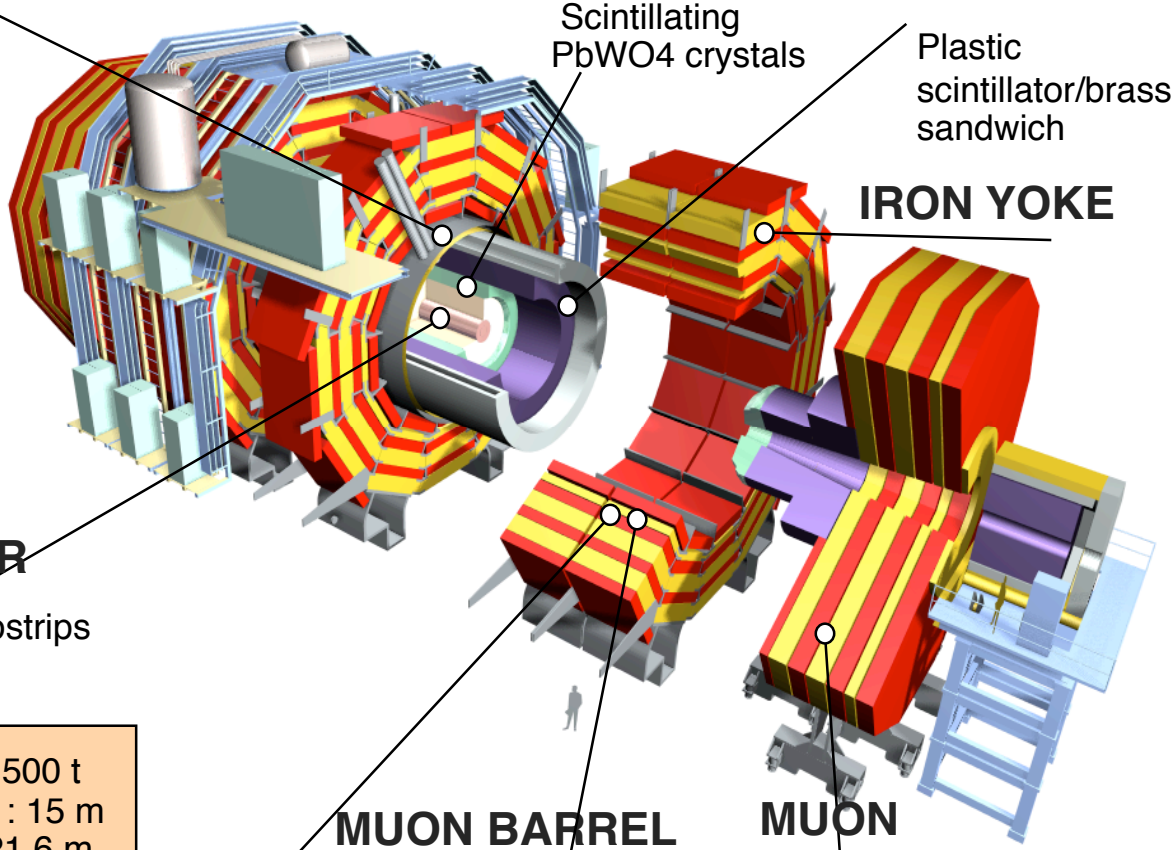
CALORIMETERS

ECAL

Scintillating PbWO4 crystals

HCAL

Plastic scintillator/brass sandwich



IRON YOKE

TRACKER

Silicon Microstrips
Pixels

Total weight : 12,500 t
Overall diameter : 15 m
Overall length : 21.6 m
Magnetic field : 4 Tesla

MUON BARREL

Drift Tube Chambers (DT)

Resistive Plate Chambers (RPC)

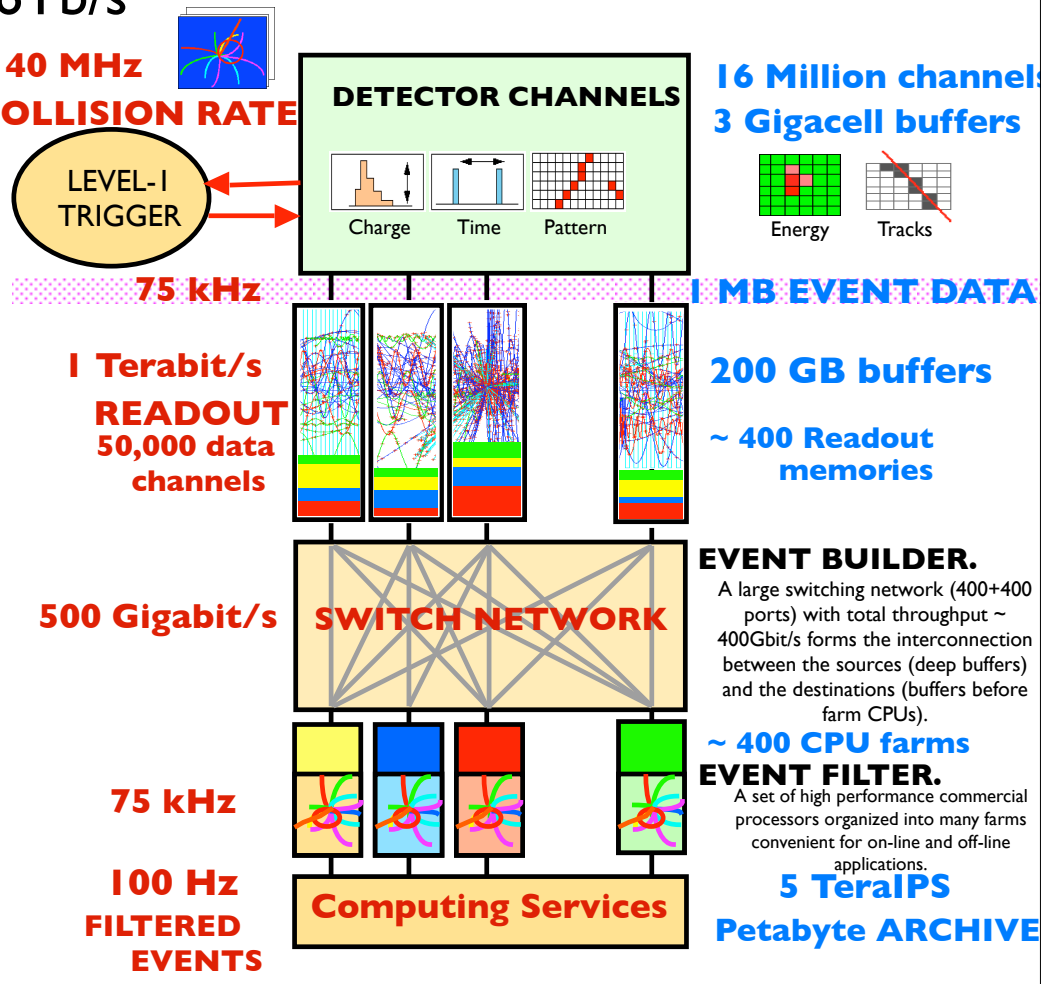
MUON ENDCAPS

Cathode Strip Chambers (CSC)

First CMS Network

The highest performance network in CMS is used before the data is selected

- The detector reads out at ~0.6Tb/s
- Feeds into 400Gb ports
- 500Gigabit/s Switch Network
  - Originally planned custom
  - Currently a composite
  - May be commodity before we start
- Filter farm selects 100Hz of events to write to tape
  - 100Hz motivated by money



Once Data is Selected

CMS events size are ~1MB and collected at 100Hz

- Roughly 1 billion events a year
- Approximately 1pB of data recorded per year
- Reconstructed data is another 0.5pB per year
- We expected up to several pB of simulated data per year
  - Generated at remote facilities

To analyze the data it is delivered to computing centers for analysis and reprocessing

- The CMS computing challenge is substantial but not a revolution over existing computing challenges from running experiments.
  - In 2007 the total computing in CMS is expected to be factors above running detectors in 2007, but not an order of magnitude

Revolutionary Aspects

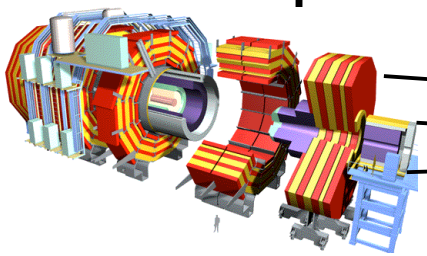
If there is a revolutionary aspect of CMS computing its the level of distribution of the computing centers

- The CMS computing model is widely distributed

For the first time the host lab is a comparatively small percentage of the total computing resources

- There will not be enough analysis resources in any one place to complete the scientific program even at the beginning
- Data Distribution, Managed Networking, and Access to Distributed Resources all has to work properly on day 1
- Distributed resources are not a way of augmenting the program but critical to the success from the beginning
- The computing for the experiment is provided by a collection of peers

CMS Distributed Computing



LHC Experiment

Online System

100-200 MBytes/s

Tier 0

CERN Computer Center > 20 TIPS

2.5 - 10 Gbits/s

Tier 1

Japan

UK

France

USA

...

2.5 Gbits/s

Tier 2

Tier2 Center

2 Center

2 Center

2 Center

~0.6 Gbits/s

Tier 3

Institute

stitute

stitute

stitute

1 Gbits/s

Physics cache

Tier 4

PCs, other portals

Several Networking Challenges

CMS is proposing replicating data from the Tier-0 to the Tier-1s in real time

- Percentage of raw data would be replicated to each Tier-1
  - Serves as a second archive copy of the data
  - Provides capability for moving rereconstruction to Tier-1 centers

Tier-1 centers which are located in 3 continents are an extension of the data acquisition system

- Network reliability and predictability are extremely important because it is difficult to recover from a significant loss of service
  - Buffers are built into the system, but nothing is allowed to slow the flow of data

Data rate is manageable

- At the start of the experiment replicating raw data is likely to require around 25MB/s
- More interesting if the experiments have the resources to double rate
- More interesting for CERN because there are 4 experiments

Analysis Network Challenge

The more traditional network problem is between the Tiers for analysis

- Tier-1s will get a copy of all reconstructed data from CERN (or other Tier-1s) for analysis
- This is replicated in part to Tier-2 centers to support analysis communities
- Simulated data is created at Tier-2 centers and archived centrally in Tier-1 mass storage systems

A Tier-1 facility has around 1pB of disk storage for analysis

- multi pB of mass storage for archiving and data serving

A Tier-2 facility has ~200TB of disk storage

- About 75% is expected for analysis
- 25% for simulation and staging space

Tier-2 Networking Estimates

Tier-2 (and subsequently Tier-1) estimates are motivated by using resources

To make flexible use of a 200TB storage facility a group needs reasonable networking

- Tier-2 centers should strive for 2.5Gb/s to 10Gb/s by the start of the experiment
  - Both Caltech, UFL, and UCSD Tier-2 prototypes either have or will shortly have this
- The networking available at the Tier-1 needs to support Tier-2 transfers

Network Speed	Time to Completely cycle Disk Storage
1Gb/s	20 days
2.5Gb/s	8 days
10Gb/s	2 days

A few Computing Model thoughts

Currently CMS model of a site is fairly traditional

- Large cache is designed for fairly static placement of data
- If sufficient networking exists it is interesting to imagine more dynamic models
- This opens new Physics Opportunities for analysis in the US

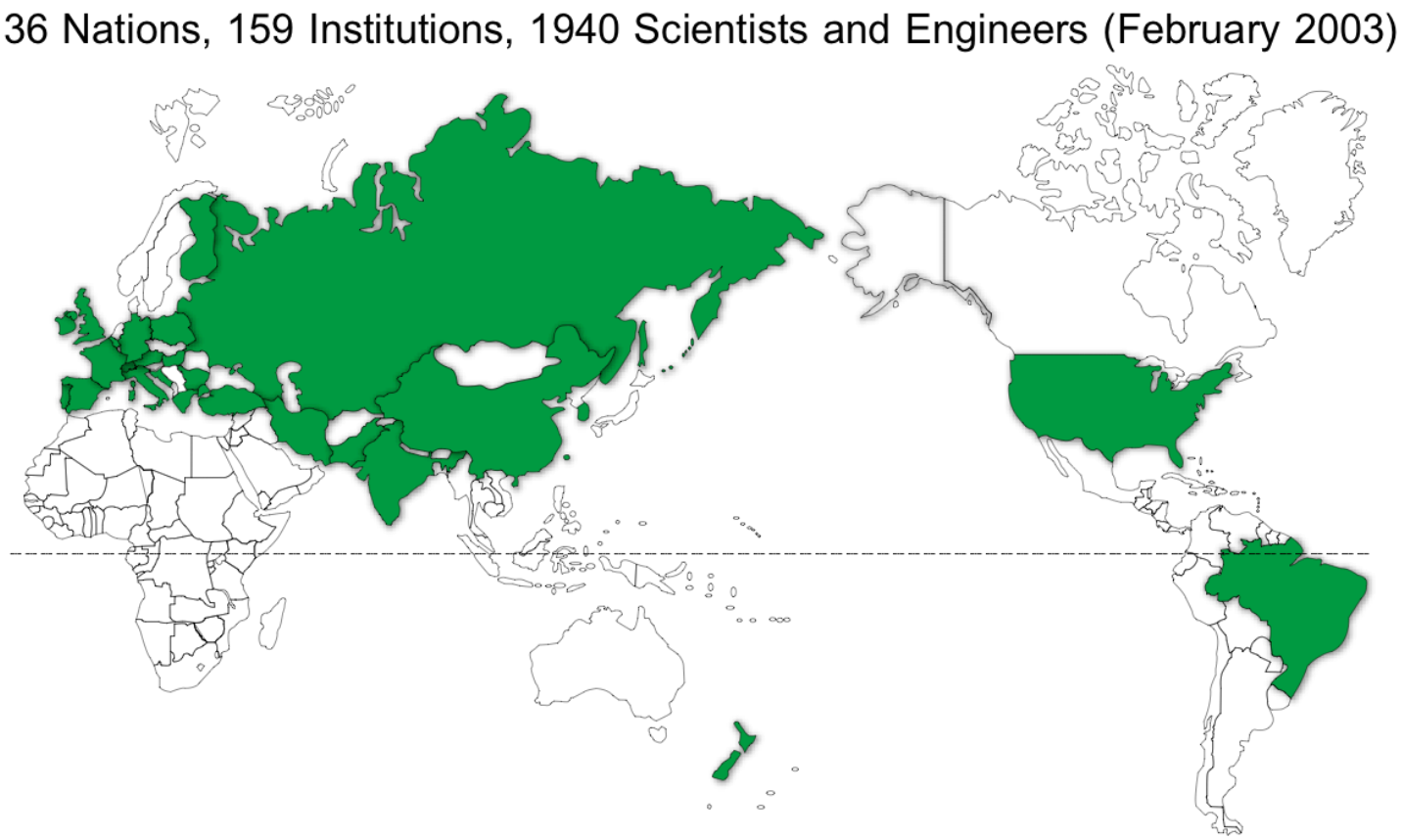
At a Tier-2 200TB of disk seems like a lot

- From an operational standpoint it is
- However, the space for analysis is only about 15% of one years raw data or about 30% of one years reconstructed data
  - Users traditionally access several years data
  - The Tier-2 program will support about 200 users (40 at an average site)
    - Potential for cache being quite dynamic

Project Structure

CMS is an international organization with a lot of players

- Plenty of managed projects to interact with



Management Structures

Contributors

- International CMS (Computing and Core Software)
  - Central coordination and technical management for the experiment
  - Includes 5 Level 2 Tasks
    - Infrastructure and Services
    - Core Application Software
      - Includes a Computing Center and a Networking Component
    - Production
    - Data Management
      - End-to-end network optimization components
    - WorkFlow Management

Management Structures (2)

National CMS Collaborations

- In the US the Software and Computing Project is
  - Hosted at Fermilab
  - Lothar Bauerdick is Project Manager
  - Project is Divided into 2 Level 2 Tasks
    - User Facilities - Ian Fisk
      - Tier-1, Tier-2 Centers and Distributed Computing Infrastructure
      - Network Issues and end-to-end optimization with CERN are here
    - Core Application Software - Bob Clare
      - Experiment Reconstruction Framework and Tools

CERN and Grid Projects

The CMS connection with CERN opens access to European Network

- CERN manages the LCG and EGEE Grid Projects
  - In Europe a lot of the facility and network issues are centrally coordinated
- CERN is the connection point for European networking